

Universally L^1 -Bad Arithmetic Sequences

Patrick LaVictoire* UC Berkeley

October 8, 2009

Abstract

We present a modified version of Buczolich and Mauldin's proof that the sequence of square numbers is universally L^1 -bad. We extend this result to a large class of sequences, including the d th powers and the set of primes; furthermore, we show that any subsequence of the averages taken along these sequences is also universally L^1 -bad.

1 Introduction

Let $(X, \mathcal{F}, \mu, \tau)$ a dynamical system. For a sequence of natural numbers $\{n_k\}$ and any $f \in L^1(X)$, we can consider the subsequence average

$$A_N f(x) := \frac{1}{N} \sum_{k=1}^N f(\tau^{n_k} x).$$

By analogy with Birkhoff's Pointwise Ergodic Theorem, we will examine the a.e. convergence or divergence of $A_N f(x)$ as $N \rightarrow \infty$.

We say that $\{n_k\}$ is *universally L^p -good* if for every dynamical system $(X, \mathcal{F}, \mu, \tau)$ and every $f \in L^p(X, \mu)$, $\lim_{N \rightarrow \infty} A_N f(x)$ exists for almost every $x \in X$. We say that $\{n_k\}$ is *universally L^p -bad* if for every non-atomic ergodic dynamical system $(X, \mathcal{F}, \mu, \tau)$, there exists an $f \in L^p(X, \mu)$ such that the sequence $\{A_N f(x)\}_{N=1}^\infty$ diverges on a set of positive measure in X . Finally, we say that $\{n_k\}$ is *persistently universally L^p -bad* if for every non-atomic ergodic dynamical system $(X, \mathcal{F}, \mu, \tau)$ and every infinite $\mathcal{S} \subset \mathbb{N}$, there exists an $f \in L^p(X, \mu)$ such that the sequence $\{A_N f(x)\}_{N \in \mathcal{S}}$ (N taken in increasing order) diverges on a set of positive measure in X .

Among the classical results in this topic, Bourgain [2] proved that (the integer part of) any sequence of polynomial values is universally L^p -good for any $p > 1$, and Bourgain [1] and Wierdl [11] showed that the same is true of the sequence of prime numbers. For these sequences, the Banach principle of Sawyer [9] implies that pointwise convergence of $A_N f$ for all $f \in L^1$ depends only on the validity of a weak maximal inequality

$$\| \sup_{N \in \mathbb{N}} |A_N f(x)| \|_{1, \infty} \leq C \|f\|_1 \quad \forall f \in L^1(X). \quad (1.1)$$

The Conze principle [6] allows the transfer of such an inequality (with the same constant) from any ergodic dynamical system $(X, \mathcal{F}, \mu, \tau)$ to any other dynamical system. Therefore, one of these sequences would be universally L^1 -good if and only if there were some fixed $C > 0$ such that (1.1) held for every dynamical system (X, \mathcal{F}, m, T) , and it would be universally L^1 -bad otherwise.

*Supported in part by NSF Grant DMS-0401260.

This question of pointwise convergence for subsequence averages of L^1 functions remained open for virtually all sequences of interest, including all polynomials of degree ≥ 2 and the sequence of primes, until Buczolich and Mauldin [3] [4] proved that $\{k^2\}$ is in fact universally L^1 -bad.

In this paper, we adapt and extend the construction in [3] to prove the following theorem:

Theorem 1.1. *Let $n_k = k^d$ for some $d > 1$, or $n_k =$ the k th prime number. Given any $C > 0$ and any infinite set $\mathcal{S} \subset \mathbb{N}$, there exists a dynamical system $(X, \mathcal{F}, \mu, \tau)$ and an $f \in L^1(X)$ such that*

$$\left\| \sup_{N \in \mathcal{S}} \left| \frac{1}{N} \sum_{k=1}^N f \circ \tau^{n_k} \right| \right\|_{1,\infty} > C \|f\|_1.$$

As discussed above, this implies

Corollary 1.2. *The sequence of d th powers ($d > 1$) and the sequence of primes are persistently universally L^1 -bad.*

This exposition is self-contained, with the exception of the number-theoretic results of Hooley [8] and Granville and Kurlberg [7].

A few words on the structure of this paper: In Section 2, we present a heuristic version of the argument, in the case of the squares. Then in Section 3, we express the general form of our result (Theorem 3.1) and prove that its conditions are indeed satisfied by the d th powers and the sequence of primes. In Section 4, we present the main inductive step (Proposition 4.1), show that it implies Theorem 3.1, and explain the structure of the induction.

In Sections 5-7, we construct the various objects of the succeeding inductive step and prove several necessary lemmas about them. Section 8 brings these parts together and proves that the properties claimed in Proposition 4.1 do indeed hold for this next step, completing the proof of Theorem 3.1. In Section 9, we retrospectively explain the purpose of several objects and lemmas in this intricate proof.

Our notation will rarely distinguish between \mathbb{Z}_N (a probability space with the measure-preserving transformation $\tau x = x + 1 \pmod{N}$) and \mathbb{Z} . Sets and functions on \mathbb{Z}_N will correspond to N -periodic sets and functions on \mathbb{Z} , and any object on \mathbb{Z}_N is understood to represent an object on \mathbb{Z}_{MN} for any $M \in \mathbb{Z}^+$.

Furthermore, we let \mathbb{P} denote the uniform probability measure on \mathbb{Z}_N , and $\mathbb{E}X$ the expected value of a random variable $X : \mathbb{Z}_N \rightarrow \mathbb{R}$. Note that the values of \mathbb{P} and \mathbb{E} are unchanged when we consider a N -periodic set or function as an object on \mathbb{Z}_{MN} instead, so that we may use \mathbb{P} and \mathbb{E} freely without keeping track of N . We will use $X \stackrel{d}{=} Y$ to denote that two random variables X and Y (not necessarily on the same probability space) have identical distributions.

Finally, we will use both subscripts and superscripts on certain functions $f_h^L, g_h^L, X_h^L, Z_h^L$ and certain sets $\Lambda_q^\gamma, \Xi_q^\gamma$. To prevent these from being confused with exponential notation, we note here that such superscripts on these objects will not denote exponents; we will therefore write the square of X_h as $(X_h)^2$ rather than X_h^2 .

2 Outline of the Argument for the Squares

Here we will present a heuristic outline of the argument in the original case $n_k = k^2$, before introducing the necessary complications (exceptional sets and the like). We therefore ask the reader's patience with these claims, some of which are not technically true; the argument presented in Section 3 and thereafter will be rigorous.

Buczolich and Mauldin's proof in [3] essentially boils down to three key insights. The first is that in order to prove that $\{k^2\}$ is universally L^1 -bad, it suffices to prove the existence of what they term (K, M) families¹ for arbitrary $K, M \in \mathbb{N}$.

Definition Given $K, M \in \mathbb{N}$ and a measure-preserving system $(\Omega, \mathcal{F}, \tau, \mathbb{P})$, a (K, M) family on Ω consists of the following:

- $f_1, \dots, f_K \in L^1(\Omega)$ with $f_h \geq 0$ and $\mathbb{E}f_h \leq 1$
- $X_1, \dots, X_K \in L^1(\Omega)$ pairwise independent with $\mathbb{E}X_h \geq M$ and $\mathbb{E}(X_h)^2 \leq C_M$
- A measurable function $Q_x : \Omega \rightarrow \mathbb{N}$ such that for a.e. $x \in \Omega$,

$$\frac{1}{Q_x} \sum_{k=1}^{Q_x} f_h(\tau^{k^2} x) \geq X_h(x) \quad \forall 1 \leq h \leq K \quad (2.1)$$

Note that for each $x \in \Omega$, Q_x does not depend on h .

The point of constructing a (K, M) family is that, while a single X_h may have a weak L^1 norm no greater than the L^1 norm of f , an average of pairwise independent random variables with uniformly bounded variance is subject to the Weak Law of Large Numbers. Thus for some large K , the average $\frac{X_1(x) + \dots + X_K(x)}{K}$ will be at least $\frac{M}{2}$ on a set of probability at least $\frac{1}{2}$, giving a weak L^1 norm of at least $\frac{M}{4}$, while the L^1 norm of the average $\frac{f_1(x) + \dots + f_K(x)}{K}$ remains ≤ 1 .

Therefore if we have (K, M) families for all $K, M \in \mathbb{N}$, we can construct a dynamical system and a function that violate any given weak (1,1) maximal inequality.

The second insight is that we can inductively construct a (K, M) family on the probability space \mathbb{Z}_N with the shift operator $\tau x = x + 1 \pmod{N}$, for some large N depending on K and M . We will use an inner induction: given a $(K-1, M)$ family, we will construct f_K and X_K in stages f_K^L, X_K^L , so that we have all the required properties except for $\mathbb{E}X_K^L \geq M$, which we replace with $\mathbb{E}X_K^{L+1} \geq \mathbb{E}X_K^L + \epsilon$. Then for L large enough, we will have a (K, M) family as desired.

The main obstruction to this approach is the difficulty of maintaining pairwise independence; when we alter f_K^L and X_K^L , we must at the same time alter f_1^L, \dots, f_{K-1}^L and X_1^L, \dots, X_{K-1}^L in order to sustain this property. We do this by taking into account properties of the distribution of the squares in residue classes.

Definition Given $\gamma > 0$, and given $q \in \mathbb{N}$ squarefree and odd with κ prime factors, consider the following subsets of the integers:

$$\begin{aligned} \Lambda_q &:= \{x \in \mathbb{Z} : (x, q) = 1, \exists k \in \mathbb{Z} \text{ such that } x \equiv k^2 \pmod{q}\} \\ (-\Lambda_q)^\gamma &:= -\Lambda_q + (0, \gamma 2^\kappa) \end{aligned}$$

It follows from elementary number theory that every $x \in \Lambda_q$ has exactly 2^κ square roots modulo q ; and that if $q = p_1 \dots p_\kappa$ with p_i large, then $\mathbb{P}(\Lambda_q) = \prod_{i=1}^\kappa \frac{p_i - 1}{2p_i} \approx 2^{-\kappa}$. Clearly $\mathbb{P}((-\Lambda_q)^\gamma) \leq \gamma$; it follows from the result of Granville and Kurlberg [7] that $\mathbb{P}((-\Lambda_q)^\gamma) \approx 1 - e^{-\gamma}$.

Having set $f_K^0 \equiv X_K^0 \equiv 1$, at each step in L we will take some highly composite $q = q_L$ and set

$$f_K^{L+1}(x) := e^\gamma f_K^L(x) \mathbf{1}_{\mathbb{Z} \setminus (-\Lambda_q)^\gamma}(x) \quad (2.2)$$

$$X_K^{L+1}(x) := e^\gamma X_K^L(x) \mathbf{1}_{\mathbb{Z} \setminus (-\Lambda_q)^\gamma}(x) + c \mathbf{1}_{\Psi_q}(x), \quad (2.3)$$

¹Buczolich and Mauldin define these families with many more conditions, in the fashion that we will define our inductive Step (K, M, L) in Proposition 4.1; but the definition here will suffice for heuristic purposes.

for some $\Psi_q \subset (-\Lambda_q)^\gamma$ with $\mathbb{P}(\Psi_q) \geq c\gamma$. To maintain the property (2.1), for $x \notin (-\Lambda_q)^\gamma$ we will keep the same length of averaging Q_x we used for f_K^L ; but for $x \in (-\Lambda_q)^\gamma$ we will redefine Q_x to be a multiple of q . We will define the set Ψ_q such that if we add a generic square number to an element of Ψ_q , we have a good chance of landing in the support of f_K^{L+1} ; the result of [7] ensures that this is possible.

We can ensure that X_K^{L+1} remains independent from our original $(K-1, M)$ family by choosing q relatively prime to the period of that family; note that $\mathbb{E}X_K^{L+1} = \mathbb{E}X_K^L + c^2$. However, we are not quite out of the woods: for $x \in (-\Lambda_q)^\gamma$ we are now using different values of Q_x with f_K^{L+1} than with our $(K-1, M)$ family.

This defect is repaired by the third insight of [3]: using the inductive hypothesis, we take a different $(K-1, M)$ family $\{g_1^L, \dots, g_{K-1}^L, Z_1^L, \dots, Z_{K-1}^L, Q'_x\}$, where the lengths of averaging Q'_x are relatively prime to Tq ; we may then “restrict” it to $(-\Lambda_q)^\gamma$ by taking

$$\begin{aligned}\bar{g}_h^L(x) &:= q2^{-\kappa} g_h^L(x) \mathbf{1}_{(0, \gamma 2^\kappa) + q\mathbb{Z}}(x) \\ \bar{Z}_h^L(x) &:= Z_h^L(x) \mathbf{1}_{(-\Lambda_q)^\gamma}(x) \\ \bar{Q}'_x &:= qQ'_x.\end{aligned}$$

This preserves the property (2.1) on $(-\Lambda_q)^\gamma$ if we assume that

$$g_h^L(\tau^{Q'_x} y) = g_h^L(y + Q'_x) = f(y) \quad \forall y \in [x, x + (qQ'_x)^2], \quad (2.4)$$

since if $x \in (-\Lambda_q)^\gamma$, then $|\{k \in [1, qQ'_x] : x + k^2 \in (0, \gamma 2^\kappa) + q\mathbb{Z}\}| \geq 2^\kappa Q'_x$ and this set is equidistributed modulo Q'_x . This implies that for $x \in (-\Lambda_q)^\gamma$,

$$\frac{1}{Q'_x} \sum_{k=1}^{Q'_x} \bar{g}_h^L(x + k^2) = \frac{1}{qQ'_x} \sum_{k=1}^{Q'_x} 2^\kappa \cdot q2^{-\kappa} g_h^L(x + k^2) \geq Z_h(x) = \bar{Z}_h(x).$$

(That is, averages of \bar{g}_h^L over long intervals of the squares look like averages of g_h^L on short intervals of the squares.) Then if for $h \leq K-1$ we let

$$\begin{aligned}f_h^{L+1}(x) &:= f_h^L(x) \mathbf{1}_{\mathbb{Z} \setminus (-\Lambda_q)^\gamma}(x) + \bar{g}_h^L \\ X_h^{L+1}(x) &:= X_h^L(x) \mathbf{1}_{\mathbb{Z} \setminus (-\Lambda_q)^\gamma}(x) + \bar{Z}_h^L\end{aligned}$$

and set the new Q_x to equal \bar{Q}'_x on the set $(-\Lambda_q)^\gamma$, we find that we have nearly preserved the properties of the family we began with. Thus we may iterate the inner inductive step (to which we must of course add a version of (2.4)) and thereby construct a (K, M) family.

3 Main Theorem

Definition For a set $\Lambda \subset \mathbb{Z}_t$, we can define

$$\begin{aligned}\mathbb{P}(\Lambda) &:= \frac{|\Lambda|}{t} \\ s_t &:= \mathbb{P}(\Lambda)^{-1} \\ \Lambda^\gamma &:= \Lambda + (0, \gamma s_t) \subset \mathbb{Z}_t.\end{aligned}$$

Remark Note that this turns \mathbb{Z}_t into a probability space, that s_t is the average spacing between elements of Λ in \mathbb{Z}_t , and that $\mathbb{P}(\Lambda^\gamma) \leq \gamma$. If $\Lambda \subset \mathbb{Z}$ is periodic by t , we can consider it as a subset of \mathbb{Z}_{nt} for any $n \in \mathbb{Z}^+$, and we see that $\mathbb{P}(\Lambda)$ is independent of n .

The main result of this paper is the following:

Theorem 3.1. Let $\{n_k\} \subset \mathbb{N}$ an increasing sequence, and $\alpha, \beta > 0$. Say that for every integer $t > 1$, there exists a set of residues $\Lambda_t \subset \{n_k + t\mathbb{Z} : k \in \mathbb{N}\} \subset \mathbb{Z}_t$ such that $\Lambda_{st} = \Lambda_s \cap \Lambda_t$ whenever $(s, t) = 1$, and that there exist some auxiliary sequences $\{p_j\}$, $\{q_j\}$ of pairwise relatively prime positive integers such that

$$\mathbb{P}(\Lambda_{q_j}) \rightarrow 0, \tag{3.1}$$

$$\inf_j \mathbb{P}(\Lambda_{p_j}) > 0, \tag{3.2}$$

$$\varepsilon_\gamma := \gamma - \liminf_{j \rightarrow \infty} \mathbb{P}(\Lambda_{q_j}^\gamma) = o(\gamma), \tag{3.3}$$

such that for all $\gamma > 0$ sufficiently small,

$$\liminf_{j \rightarrow \infty} |\{(u, v) \in \Lambda_{q_j} \times \Lambda_{q_j} : |u - v - w| > \gamma s_{q_j} \ \forall w \in \Lambda_{q_j}\}| > 5\alpha |\Lambda_{q_j}|^2, \tag{3.4}$$

and for all $Q = p_{i_1} \dots p_{i_k} q_{j_1} \dots q_{j_l}$ with $i_1 < \dots < i_k, j_1 < \dots < j_l$,

$$\liminf_{N \rightarrow \infty} \frac{1}{N} |\{1 \leq k \leq N : n_k \equiv a \pmod{Q}\}| > \frac{\beta}{|\Lambda_Q|} \quad \forall a \in \Lambda_Q. \tag{3.5}$$

Then, given any $C > 0$ and any infinite set $\mathcal{S} \subset \mathbb{N}$, there exists a probability space $(X, \mathcal{F}, \mathbb{P})$, a measure-preserving transformation τ on X , and an $f \in L^1(X)$ such that

$$\left\| \sup_{N \in \mathcal{S}} \left| \frac{1}{N} \sum_{k=1}^N f \circ \tau^{n_k} \right| \right\|_{1, \infty} > C \|f\|_1.$$

Remark (3.3) states that Λ_{q_j} does not cluster too much in \mathbb{Z}_{q_j} so that $\mathbb{P}(\Lambda_{q_j}^\gamma)$ is nearly γ , while (3.4) states that the set of differences of elements in Λ_{q_j} is not concentrated near Λ_{q_j} . (3.5) states that each point of Λ_Q is hit uniformly often by the sequence $\{n_k\}$, if Q is any squarefree product of terms from the auxiliary sequences $\{p_j\}$ and $\{q_j\}$. (The function $x \rightarrow Q_x$ we will later construct will take such products as its values.)

Claim. Theorem 3.1 implies Theorem 1.1.

Proof. For the sequence of primes, we take Λ_t to be the set of integers relatively prime to t , then let p_j be distinct primes and q_j be highly composite such that $q_j^{-1} \phi(q_j) \rightarrow 0$; since $\prod_{p \text{ prime}} \frac{p-1}{p} = 0$, it is possible to choose such q_j . The property (3.5) is clear from Dirichlet's theorem on arithmetic progressions.

For the sequence of d th powers, we will take Λ_t to be the residues of $k^d \pmod{t}$ which are units mod t . We will take p_j to be distinct primes congruent to 1 mod d , and q_j to be products of j such primes. Note that if Q has κ prime factors each congruent to 1 mod d , then every $x \in \Lambda_Q$ will have precisely d^κ d th roots in \mathbb{Z}_Q , and $\mathbb{P}(\Lambda_Q) = \phi(Q)d^{-j}$. If we choose all of the prime factors sufficiently large, we can ensure that $\phi(Q) \geq \frac{1}{2}$ for all squarefree products Q of these sequences, so that (3.5) is satisfied.

To prove (3.3) and (3.4) for each of these cases, we will use some recent results on the distribution of the residues Λ_{q_j} (as the average spacing $s_{q_j} \rightarrow \infty$); roughly speaking, in each case they are distributed locally like a Poisson process of rate $\mathbb{P}(\Lambda_{q_j})$. We begin by introducing some notation.

Definition Let $\{y_k\}_{k=1}^N$ be a strictly increasing sequence of real numbers in $[0, N]$, and consider it as a subset of $\mathbb{T} = \mathbb{R}/N\mathbb{Z}$. For $E \subset \{1, \dots, N\}$, and $\theta > 0$, we define a probability measure and a cumulative distribution function:

$$\tilde{\mathbb{P}}_k(E) := \frac{|E|}{N} \tag{3.6}$$

$$F(\theta) := \tilde{\mathbb{P}}_k(|y_{k+1} - y_k| > \theta). \tag{3.7}$$

Now if we consider the normalized set $s_q^{-1}\Lambda_q = \{y_i : 1 \leq i \leq |\Lambda_q|\} \subset [0, |\Lambda_q|]$ (where the y_i are taken in increasing order), we can examine the cumulative distribution function $F_q(\theta)$ defined as above. Hooley [8] proves for the primes, and Granville and Kurlberg [7] prove for the d th powers², that $F_q(\theta) \rightarrow e^{-\theta}$ pointwise as $s_q \rightarrow \infty$ (for which reason they call these sets ‘‘Poisson distributed’’). Thus

$$\mathbb{P}(\Lambda_q) = |\Lambda_q|^{-1} \sum_i |y_{i+1} - y_i| \wedge \gamma = \int_0^\gamma F_q(\theta) d\theta \rightarrow 1 - e^{-\gamma},$$

so that we have (3.3). To prove (3.4) from this distributional fact, however, requires a little more work.

Lemma 3.2. *Let $\zeta_l(\theta) := \tilde{\mathbb{P}}_i(|y_i - y_k - y_l| > \theta \forall k)$. For $\theta > 0$ and $J > 4$,*

$$\tilde{\mathbb{P}}_l \left(\zeta_l(\theta) > \frac{F(J\theta) + F(2\theta) - 1}{2} \right) \geq F(2\theta) - F((J-2)\theta).$$

Proof. Define $A := \{k : |y_{k+1} - y_k| \leq J\theta\}$; then $\tilde{\mathbb{P}}_k(A) = 1 - F(J\theta)$.

For $z \in [0, N]$, let

$$\begin{aligned} B_z &:= \{i : \exists k \in A : |y_i - y_k - z| \leq \theta\} \\ C_z &:= \{i : \exists k \notin A : |y_i - y_k - z| \leq \theta\}. \end{aligned}$$

There are $|B_z|$ points $y_i - z$ contained in the $|A|$ intervals $[y_k - \theta, y_k + \theta]$; thus it must be that for at least $|B_z| - |A|$ of these points, their successors $y_{i+1} - z$ also lie in the same interval, which implies that $|y_{i+1} - y_i| \leq 2\theta$. Thus $\tilde{\mathbb{P}}_i(B_z) \leq \tilde{\mathbb{P}}_k(A) + 1 - F(2\theta) = 2 - F(J\theta) - F(2\theta)$.

Therefore, for any l ,

$$\begin{aligned} 1 - \zeta_l(\theta) &\leq \tilde{\mathbb{P}}_i(\exists k \notin A : |y_i - y_k - y_l| \leq \theta) + \tilde{\mathbb{P}}_i(\exists k \in A : |y_i - y_k - y_l| \leq \theta) \\ &= \tilde{\mathbb{P}}_i(C_{y_l}) + \tilde{\mathbb{P}}_i(B_{y_l}) \\ &\leq \tilde{\mathbb{P}}_i(C_{y_l}) + 2 - F(J\theta) - F(2\theta). \end{aligned}$$

But if $i \in C_{y_l}$ and $2\theta < |y_l - y_{l-1}| < (J-2)\theta$, then

$$y_i - y_{l-1} \in (y_k + \theta, y_{k+1} - \theta)$$

so that $|y_i - y_k - y_{l-1}| > \theta \forall k$. This implies that

$$\zeta_l(\theta) + \zeta_{l-1}(\theta) \geq \zeta_l(\theta) + \tilde{\mathbb{P}}_i(C_{y_l}) \geq F(J\theta) + F(2\theta) - 1$$

whenever $y_l - y_{l-1} \in (2\theta, (J-2)\theta)$.

Therefore

$$\begin{aligned} \tilde{\mathbb{P}}_l \left(\zeta_l(\theta) > \frac{F(J\theta) + F(2\theta) - 1}{2} \right) &\geq \frac{1}{2} \tilde{\mathbb{P}}_l(\zeta_l(\theta) + \zeta_{l-1}(\theta) > F(J\theta) + F(2\theta) - 1) \\ &\geq \tilde{\mathbb{P}}_l(2\theta < y_l - y_{l-1} \leq (J-2)\theta) \\ &\geq F(2\theta) - F((J-2)\theta). \end{aligned}$$

□

²While Corollary 2 in that paper is stated for the entire set of d th power residues modulo q , we see that passing to the subset Λ_q does not change equation (1) and thus we may apply the same result.

Now if $\gamma < \frac{1}{12}$, then $e^{-2\gamma} > \frac{6}{7}$ and we can set $J := \frac{\log 2}{\gamma} > 4$; then Lemma 3.2 implies

$$\begin{aligned}\tilde{\mathbb{P}}_l(\zeta_l > 1/8) &\geq \tilde{\mathbb{P}}_l\left(\zeta_l > \frac{-1 + e^{-J\gamma} + e^{-2\gamma}}{2} + o_j(1)\right) \\ &\geq e^{-2\gamma} - e^{(2-J)\gamma} + o_j(1) \geq \frac{1}{4}\end{aligned}$$

for j sufficiently large, so that (3.4) holds with $\alpha = \frac{1}{160}$. \square

Remark Note that (3.5) cannot be satisfied by polynomials other than $n_k = c_d k^d + c_0$, as can be seen from Cohen's result [5] that if we fix $a \in \mathbb{Z}_p[x]$ of degree d which is not of this type and consider $a(x) - y$ as $y \in \mathbb{Z}_p$ varies, for some fixed proportion of y this polynomial will have d distinct roots, while for some fixed proportion it will have 1 root. This prevents (3.5) from holding for sufficiently composite products. However, the author expects that every polynomial of degree 2 or greater over \mathbb{Z} should be persistently universally L^1 -bad, and that some clever variant of this argument should suffice to prove as much.

4 The Inductive Step

For an inductive argument to work, we will have to specify additional properties of the objects we seek, including a fixed distribution for the functions X_h^L .

Definition Given the sequences $\{p_j\}$ and $\{q_j\}$ as in Theorem 3.1, denote the set of their (squarefree) products

$$\mathcal{Q} := \{p_{i_1} \dots p_{i_k} q_{j_1} \dots q_{j_l} : i_1 < \dots < i_k, j_1 < \dots < j_l\} \quad (4.1)$$

and the functions (depending on the infinite $\mathcal{S} \subset \mathbb{N}$ in Theorem 3.1)

$$N(Q) := \max \left\{ N : \exists a \in \Lambda_Q : |\{1 \leq k \leq N : n_k \equiv a \pmod{Q}\}| \leq \frac{\beta N}{|\Lambda_Q|} \right\} \quad (4.2)$$

$$\psi(n) := \inf\{s \in \mathcal{S} : s > N(Q) \ \forall Q \in \mathcal{Q}, Q \leq n\}. \quad (4.3)$$

Definition For $0 < \gamma < 1$ and $0 < \alpha < 1$, we recursively define functions $Y_{n,\gamma,\alpha} : [0, 1]^{n+1} \rightarrow \mathbb{R}$, which we may consider as random variables with the Lebesgue measure. Let $Y_{0,\gamma,\alpha}(x_0) \equiv 1$, and

$$Y_{n+1,\gamma,\alpha}(x_0, \dots, x_{n+1}) := (1 - \gamma)^{-1} Y_{n,\gamma,\alpha}(x_0, \dots, x_n) \mathbf{1}_{[\gamma,1)}(x_{n+1}) + \alpha \mathbf{1}_{[0,\alpha\gamma)}(x_{n+1}). \quad (4.4)$$

Note that

$$\begin{aligned}\mathbb{E}(Y_{n,\gamma,\alpha}) &= \mathbb{E}Y_{n-1,\gamma,\alpha} + \alpha^2 \gamma = 1 + n\alpha^2 \gamma, \\ \mathbb{E}(Y_{n,\gamma,\alpha})^2 &= (1 - \gamma)^{-1} \mathbb{E}(Y_{n-1,\gamma,\alpha})^2 + \alpha^3 \gamma \leq (1 - \gamma)^{-n} (1 + \alpha^3) \leq 2(1 - \gamma)^{-n}.\end{aligned}$$

Given the conditions of Theorem 3.1, we may assume that all the p_j and q_j are odd and that α is a dyadic rational. Choose $\gamma_0 < 1/2$ such that for all $0 < \gamma < \gamma_0$, (3.3) and (3.4) hold and $\varepsilon_\gamma < \alpha\gamma$; we will write $\varepsilon = \varepsilon_\gamma$ from now on unless otherwise specified. We are now ready to state the inductive step:

Proposition 4.1 (Step (K, M, L)). *Assume the conditions of Theorem 3.1. Given any dyadic rational $0 < \gamma < \gamma_0$, any $K, M, L \in \mathbb{N}$ with $L \leq M$, constants $A, \delta > 0$ with $\delta < 8\varepsilon$, and an odd integer D , there exist $T \in \mathcal{Q}$ and $R \in \mathbb{N}$ with $(T, D) = (R, D) = 1$ and the following objects:*

(1) $f_1, \dots, f_K \in \ell^1(\mathbb{Z}_T)$ with $f_h \geq 0$ and $1 \leq \mathbb{E}f_h \leq (1 + 4\varepsilon)^{(K-1)M+L}$.

(2) $X_1, \dots, X_K \in \ell^1(\mathbb{Z}_{RT})$ pairwise independent with $X_h \stackrel{d}{=} Y_{M,\gamma,\alpha}$ $\forall h < K$ and $X_K \stackrel{d}{=} Y_{L,\gamma,\alpha}$

(3) An exceptional set $E \subset \mathbb{Z}_T$ with $\mathbb{P}(E) \leq \delta$

(4) $Q_x : \mathbb{Z}_T \rightarrow \mathbb{Z}^+$ with $Q_x \in \mathcal{Q}$ and $Q_x \mid T \forall x$, such that for each $x \notin E$,

$$\frac{1}{|\Lambda_{Q_x}|} \sum_{\substack{a \in \Lambda_{Q_x}, \\ 1 \leq a \leq Q_x}} f_h(x+a) \geq X_h(x) \quad \forall 1 \leq h \leq K \quad (4.5)$$

and

$$f_h(x+y-Q_x) = f_h(x+y) \quad \forall 1 \leq h \leq K, Q_x \leq y \leq \psi(AQ_x). \quad (4.6)$$

Remark In the final section of this paper, we will discuss the significance of the parameters A, D, R , and δ , as well as the distribution $Y_{n,\gamma,\alpha}$, the reason we require γ and α to be dyadic rationals, and other points whose necessity in the argument is not immediately obvious. For the time being, we ask the reader's trust that these complications are required in order to make a strong enough inductive step.

Claim. *Proposition 4.1 implies Theorem 3.1.*

Proof. If we fix $C > 0$, take $\gamma > 0$ small, and set $M = \lfloor C/\gamma \rfloor$ and $K = \lfloor \gamma/C\varepsilon \rfloor$, and take $\delta \leq \frac{1}{4}$, $A = D = 1$, then at Step (K, M, M) we have

$$\begin{aligned} \mathbb{E}f_h &\leq (1+4\varepsilon)^{KM} \leq (1+4\varepsilon)^{1/\varepsilon} \leq e^4, \\ \mathbb{E}X_h &= \mathbb{E}Y_{M,\gamma,\alpha} \geq 1 + M\alpha^2\gamma \geq C\alpha^2, \\ \mathbb{E}(X_h)^2 &= \mathbb{E}Y_{M,\gamma,\alpha}^2 \leq 2(1-\gamma)^{-M} \leq 2e^{2\gamma M} \leq 2e^{2C}. \end{aligned}$$

Therefore, since the X_h are pairwise independent and identically distributed,

$$\begin{aligned} \mathbb{P}\left(\frac{X_1 + \dots + X_K}{K} \leq C\alpha^2/2\right) &\leq \mathbb{P}\left(\left|\frac{X_1 + \dots + X_K}{K} - \mathbb{E}X_1\right| \geq C\alpha^2/2\right) \\ &\leq (C\alpha^2/2)^{-2} \mathbb{E}\left|\frac{X_1 + \dots + X_K}{K} - \mathbb{E}X_1\right|^2 \\ &\leq 4\alpha^{-2}C^{-2}K^{-1}\mathbb{E}(X_1 - \mathbb{E}X_1)^2 \\ &\leq \frac{8e^{2C}}{\alpha^2 C^2 K} \end{aligned}$$

and for γ sufficiently small (by (3.3), this means $K = \lfloor \gamma/C\varepsilon \rfloor$ sufficiently large), this is less than $\frac{1}{2}$.

Now if we consider the probability space \mathbb{Z}_{RT} with the measure-preserving transformation $\tau x = x+1$, and set $f = f_1 + \dots + f_K$ and $X = X_1 + \dots + X_K$, then by (3.5), (4.5) and (4.6)

$$\frac{1}{\psi(Q_x)} \sum_{k=1}^{\psi(Q_x)} f(x+n_k) > \frac{\beta}{|\Lambda_{Q_x}|} \sum_{a \in \Lambda_{Q_x}} f(x+a) \geq \beta X(x) \quad \forall x \notin E, \quad (4.7)$$

and since $\psi(Q_x) \in \mathcal{S}$, this implies

$$\mathbb{P}\left(\sup_{N \in \mathcal{S}} \frac{1}{N} \sum_{k=1}^N f \circ \tau^{n_k} > CK\alpha^2\beta/2\right) \geq \mathbb{P}\left(\frac{X}{K} > \frac{C\alpha^2}{2}\right) - \mathbb{P}(E) \geq \frac{1}{4}.$$

Therefore

$$\left\| \sup_{N \in \mathcal{S}} \frac{1}{N} \sum_{k=1}^N f \circ \tau^{n_k} \right\|_{1,\infty} \geq \frac{CK\alpha^2\beta}{8} \geq \frac{C\alpha^2\beta}{8e^4} \|f\|_1.$$

Since C is arbitrary, there can be no maximal inequality. \square

We will prove Proposition 4.1 by induction on K and L . We fix M and γ at the beginning of the argument (since they will not change as K and L change), and prove each Step (K, M, L) for all values of A, δ and D .

Note that Step $(1, M, 0)$ is trivial, and that Step $(K-1, M, M)$ implies Step $(K, M, 0)$: fix the parameters A, δ and D and obtain $\{f_1, \dots, f_{K-1}, X_1, \dots, X_{K-1}, E, Q_x\}$ satisfying (1)-(4) on \mathbb{Z}_T . Now set $f_K \equiv 1$, $X_K \equiv 1$ on \mathbb{Z} . This clearly satisfies the conditions.

Therefore, to prove Theorem 3.1, it suffices to show that if $L < M$ and we know Step (K, M, L) and all previous steps (for all values of A, δ and D , and for a fixed γ), we can prove Step $(K, M, L+1)$ for a fixed A, δ and D and the same γ . We start by applying Step (K, M, L) with $A, \delta/4$ and D to obtain a family $\{T_L, R_L, f_1^L, \dots, f_K^L, X_1^L, \dots, X_K^L, E_L, Q_{x,L}\}$ satisfying (1)-(4); we must then construct $\{T_{L+1}, R_{L+1}, f_1^{L+1}, \dots, f_K^{L+1}, X_1^{L+1}, \dots, X_K^{L+1}, E_{L+1}, Q_{x,L+1}\}$ to have the required properties.

5 Periodic Rearrangements

As in [3], we will essentially construct the next functions by choosing a new q and redefining the current functions on the subset $(-\Lambda_q)^\gamma$. On this set, we will be selecting a new $Q_{x,L+1} \gg T_L$, and we want to ensure that the left side of (4.5) will be uniformly large on a significant subset of $(-\Lambda_q)^\gamma$. Since an average of f_h^L over Λ_Q (in the sense of (4.5) may be irregular for large Q , we will modify the f_h^L in advance so that these averages will be bounded below by a constant, while preserving their averages over Λ_Q for smaller Q .

Following Buczolich and Mauldin, we call this modification a *periodic rearrangement*. Given natural numbers $p \gg T$ with $(p, T) = 1$, we will define a linear operator $f \rightarrow \tilde{f}$ from $\ell^1(\mathbb{Z}_T)$ to $\ell^1(\mathbb{Z}_{pT})$ which preserves joint distribution of functions, such that on long blocks each \tilde{f} is identical to a translate of f .

In our particular cases, where Λ_T consists of the residues of d th powers or the integers relatively prime to T , we can simply define

$$\tilde{f}(x) := \begin{cases} f(y) & x \in y + p\mathbb{Z}, 0 \leq y < T \cdot \lfloor p/T \rfloor \\ f(x) & \text{otherwise,} \end{cases}$$

and prove directly that for every $x \in \mathbb{Z}_{pT}$,

$$\frac{1}{|\Lambda_{pT}|} \sum_{a \in \Lambda_{pT}} \tilde{f}(x+a) \geq \frac{1}{2} \mathbb{E}f.$$

(For the former case, we would use the Pólya-Vinogradov inequality on character sums; for the latter we would use Dirichlet's theorem.)

However, we lack such tools when considering more general sequences, so we shall instead use an external randomization in our construction of \tilde{f} . Let Ω be a probability space and $\xi_i(\omega)$ be independent random variables on Ω , each with a uniform distribution on the discrete set $\{0, \dots, T-1\}$. Then for any $f \in \ell^1(\mathbb{Z}_T)$ and $\omega \in \Omega$ we define (on the interval $[0, pT)$) the function

$$\tilde{f}^\omega(x) := \begin{cases} f(x + \xi_i(\omega)) & (i-1)\lfloor \sqrt{p} \rfloor T \leq x \leq i\lfloor \sqrt{p} \rfloor T, 1 \leq i \leq \sqrt{p}; \\ f(x) & (\lfloor \sqrt{p} \rfloor)^2 T \leq x < pT. \end{cases} \quad (5.1)$$

Heuristically speaking, we break \mathbb{Z}_{pT} into blocks of size $T\lfloor \sqrt{p} \rfloor$ and shift each by a random variable $\xi_i(\omega)$. The point of this is that, although we cannot directly prove that Λ_{pT} will be equidistributed in the residue classes modulo T , there must exist some value of these shifts under which this is approximately so.

Similarly, for a set $E \subset \mathbb{Z}_T$ we can define $\tilde{E}^\omega := \text{supp } \tilde{1}_E^\omega \subset \mathbb{Z}_{pT}$. Note that \tilde{E}^ω contains exactly $\lfloor \sqrt{p} \rfloor |E|$

points in each full block, and $(p - (\lfloor \sqrt{p} \rfloor)^2)|E|$ points in the last block; thus we see that $\mathbb{P}(\tilde{E}^\omega) = \mathbb{P}(E)$ for all $E \subset \mathbb{Z}_T$ and all $\omega \in \Omega$, from which it follows that this periodic rearrangement preserves the joint distribution of any collection of functions.

For most $x \in \mathbb{Z}_{pT}$ there exists an \tilde{x} such that $\tilde{f}^\omega(x + z) = f(\tilde{x} + z)$ for $0 \leq z \ll \sqrt{p}$. In particular, if we take our original exceptional set $E \subset \mathbb{Z}_T$, we can define a new exceptional set

$$E^1(\omega) := \tilde{E}^\omega \cup \bigcup_{0 \leq i \leq \frac{\sqrt{p}}{T} - 1} [iT\lfloor \sqrt{p} \rfloor - \psi(AT), iT\lfloor \sqrt{p} \rfloor] + p\mathbb{Z} \quad (5.2)$$

such that (4.5) and (4.6) are still satisfied for $\tilde{f}_h^{L,\omega}$, $\tilde{X}_h^{L,\omega}$ and $\tilde{Q}_{x,L}^\omega$ off of $E^1(\omega)$. Furthermore,

$$\mathbb{P}(E^1(\omega)) \leq \mathbb{P}(E) + \frac{\psi(AT) + 2T}{T\sqrt{p}} \leq \delta/2$$

for p sufficiently large.

Now we can prove the following lemma:

Lemma 5.1. *For p sufficiently large with $(p, T) = 1$, there exists $\omega \in \Omega$ such that for all $0 \leq f \in \ell^1(\mathbb{Z}_T)$ and all $x \in \mathbb{Z}_{pT}$,*

$$\frac{1}{|\Lambda_{pT}|} \sum_{a \in \Lambda_{pT}} \tilde{f}^\omega(x + a) \geq \frac{1}{2} \mathbb{E}f.$$

Proof. By the linearity of the periodic rearrangement, it suffices to prove that for some ω , this holds for all characteristic functions of singletons in \mathbb{Z}_T : thus it is enough to show that for all $x \in \mathbb{Z}_{pT}$ and $b \in \mathbb{Z}_T$,

$$\mathbb{P}_\omega \left(\sum_{a \in \Lambda_{pT}} \tilde{1}_{\{b\}}^\omega(x + a) < \frac{|\Lambda_{pT}|}{2T} \right) < \frac{1}{pT^2}.$$

Fix b and x ; then counting only the main blocks,

$$\begin{aligned} \sum_{a \in \Lambda_{pT}} \tilde{1}_{\{b\}}^\omega(x + a) &\geq \sum_{1 \leq i \leq \sqrt{p}} |\{a \in \Lambda_{pT} : (i-1)\lfloor \sqrt{p} \rfloor T \leq x + a < i\lfloor \sqrt{p} \rfloor T, x + a + \xi_i(\omega) - b \in T\mathbb{Z}\}| \\ &= \sum_i |\Lambda_{pT} \cap P_i(\omega)|, \end{aligned}$$

where $P_i(\omega) := \{(i-1)\lfloor \sqrt{p} \rfloor T - x, i\lfloor \sqrt{p} \rfloor T - x) \cap \{b - x - \xi_i + T\mathbb{Z}\}\}$ is an arithmetic progression in \mathbb{Z}_{pT} . If we then define

$$\nu_i(\omega) := p^{-1/2} |\Lambda_{pT} \cap P_i(\omega)|,$$

we see that $0 \leq \nu_i \leq 1$, that the ν_i are independent, and that each $z \in \Lambda_{pT} \cap [(i-1)\lfloor \sqrt{p} \rfloor T - x, i\lfloor \sqrt{p} \rfloor T - x)$ contributes to $\nu_i(\omega)$ for precisely one value of $\xi_i(\omega)$. Therefore

$$\begin{aligned} \mathbb{E}_\omega \nu_i(\omega) &= \frac{1}{T\sqrt{p}} |\Lambda_{pT} \cap [(i-1)\lfloor \sqrt{p} \rfloor T - x, i\lfloor \sqrt{p} \rfloor T - x)|, \\ \mathbf{Var}_\omega \nu_i(\omega) &\leq \mathbb{E}_\omega \nu_i^2(\omega) \leq \mathbb{E}_\omega \nu_i(\omega). \end{aligned}$$

Recall that $\inf_j \mathbb{P}(\Lambda_{p_j}) > 0$, so that for $p \gg T$, $|\Lambda_{pT}| = |\Lambda_p| \cdot |\Lambda_T| \gg T\sqrt{p}$. Thus for p sufficiently large,

$$\begin{aligned} \mathbb{E}_\omega \left(\sum_i \nu_i(\omega) \right) &\geq \frac{1}{T\sqrt{p}} |\Lambda_{pT} \setminus [x - T(p - (\lfloor \sqrt{p} \rfloor)^2), x]| > \frac{3|\Lambda_{pT}|}{4T\sqrt{p}} \\ \mathbf{Var}_\omega \left(\sum_i \nu_i(\omega) \right) &\leq \frac{|\Lambda_{pT}|}{T\sqrt{p}} \end{aligned}$$

and we may apply Chernoff's Inequality (Theorem 1.8 from [10]) to find

$$\begin{aligned}
\mathbb{P}_\omega \left(\sum_{a \in \Lambda_{pT}} \tilde{1}_{\{b\}}^\omega(x+a) < \frac{|\Lambda_{pT}|}{2T} \right) &\leq \mathbb{P}_\omega \left(\sum_i |\Lambda_{pT} \cap P_i(\omega)| < \frac{|\Lambda_{pT}|}{2T} \right) \\
&= \mathbb{P}_\omega \left(\sum_i \nu_i(\omega) < \frac{|\Lambda_{pT}|}{2T\sqrt{p}} \right) \\
&\leq \mathbb{P}_\omega \left(\sum_i \nu_i(\omega) - \mathbb{E}_\omega \nu_i > \frac{|\Lambda_{pT}|}{4T\sqrt{p}} \right) \\
&\leq 2 \exp\left(-\frac{|\Lambda_{pT}|}{64T\sqrt{p}}\right) < \frac{1}{pT^2}
\end{aligned}$$

for p sufficiently large (depending on T). \square

6 Defining f_K^{L+1}

Using the properties (3.1)-(3.4), we now take $p = p_{i(K,L)}$ and $q = q_{j(K,L)}$ from our auxiliary sequences such that they are relatively prime to each other and to T_L , R_L and D , such that p is large enough for Lemma 5.1 and such that $q \gg D$, $s_q > 8\delta^{-1}\psi(AT)$ and

$$\begin{aligned}
\mathbb{P}((- \Lambda_q)^\gamma) &= \mathbb{P}(\Lambda_q^\gamma) \geq \gamma - \varepsilon, \\
|\{(u, v) \in \Lambda_q \times \Lambda_q : |u - v - w| > \gamma s_q \ \forall w \in \Lambda_q\}| &\geq 5\alpha|\Lambda_q|^2.
\end{aligned}$$

We would like to define f_K^{L+1} and X_K^{L+1} as in (2.2) and (2.3), but then X_K^{L+1} will not precisely equal $Y_{L+1,\gamma,\alpha}$ in distribution. This is the reason we will make the X_h^{L+1} periodic by $R_{L+1}T_{L+1}$ rather than just T_{L+1} : we will later multiply the parts of X_K^{L+1} by the characteristic function of intervals whose lengths are appropriate multiples of T_{L+1} . It will be essential (for its use in later inductive steps) that we keep $(R_{L+1}, D) = 1$, and for this we will need to define f_K^{L+1} and X_K^{L+1} in a more complicated fashion.

First, we will let

$$\Phi_q := \{u \in -\Lambda_q : |\{v \in \Lambda_q : u + v \notin -\Lambda_q + (-\gamma s_q, \gamma s_q)\}| \geq 2\alpha|\Lambda_q|\} \quad (6.1)$$

and note that $|\Phi_q| \geq 3\alpha|\Lambda_q|$. We then consider $\Phi_q^\gamma = \Phi_q + (0, s_q\gamma) \subset (-\Lambda_q)^\gamma$. Now $\mathbb{P}(\Phi_q^\gamma) \geq 2\alpha\gamma$ since

$$\gamma - \varepsilon \leq \mathbb{P}((- \Lambda_q)^\gamma) \leq \mathbb{P}(\Phi_q^\gamma) + q^{-1}\gamma s_q | -\Lambda_q \setminus \Phi_q | \leq \mathbb{P}(\Phi_q^\gamma) + \gamma(1 - 3\alpha)$$

and we have stipulated that $\varepsilon \leq \alpha\gamma$.

Now we choose two sets $\Psi_q, \Delta_q \subset \mathbb{Z}_q$ such that

$$\Psi_q \subset \Phi_q^\gamma \quad (6.2)$$

$$\mathbb{P}(\Psi_q) \geq \alpha\gamma \quad (6.3)$$

$$\mathbb{Z}_q \setminus (-\Lambda_q)^\gamma \subset \Delta_q \subset \mathbb{Z}_q \setminus \Phi_q \quad (6.4)$$

$$\mathbb{P}(\Delta_q) < 1 - \mathbb{P}(\Lambda_q^\gamma) + \delta/8 \quad (6.5)$$

$$(|\Psi_q|, D) = (|\Delta_q|, D) = 1. \quad (6.6)$$

(As $q \gg D$, this last condition is clearly possible to satisfy simultaneously with the others.) We observe that

$$|\{v \in \Lambda_q : x + v \in \Delta_q\}| \geq 2\alpha|\Lambda_q| \quad \forall x \in \Psi_q, \quad (6.7)$$

and accordingly we define

$$f_K^{L+1} := (1 - \gamma)^{-1} \tilde{f}_K^L \mathbf{1}_{\Delta_q} \in \ell^1(\mathbb{Z}_{qpT}).$$

Note that $(q, pT) = 1$ implies $\mathbb{E} f_K^{L+1} = (1 - \gamma)^{-1} \mathbb{P}(\Delta_q) \mathbb{E} \tilde{f}_K^L = (1 - \gamma)^{-1} \mathbb{P}(\Delta_q) \mathbb{E} f_K^L$; since $\gamma < 1/2$ and $\delta < 8\epsilon$, this means

$$\mathbb{E} f_K^L \leq \mathbb{E} f_K^{L+1} \leq (1 + 4\epsilon) \mathbb{E} f_K^L. \quad (6.8)$$

The goal of Lemma 5.1 and the definition of Ψ_q is the following lemma:

Lemma 6.1. *Let $T = T_L$, with p and q chosen as above. For all $x \in \Psi_q$ and for any $Q \in \mathcal{Q}$ such that $qpT \mid Q$,*

$$\frac{1}{|\Lambda_Q|} \sum_{a \in \Lambda_Q} f_K^{L+1}(x + a) \geq \alpha.$$

Proof. Let $B := \frac{Q}{qpT}$. Since $Q \in \mathcal{Q}$ is squarefree, we see that B, T, p and q are pairwise relatively prime, and

$$\begin{aligned} \frac{1}{|\Lambda_Q|} \sum_{a \in \Lambda_Q} f_K^{L+1}(x + a) &= \frac{1}{|\Lambda_B|} \sum_{w \in \Lambda_B} \frac{1}{|\Lambda_{qpT}|} \sum_{z \in \Lambda_{qpT}} f_K^{L+1}(x + z) \\ &= \frac{1}{|\Lambda_{pT}|} \frac{1}{|\Lambda_q|} \sum_{u \in \Lambda_{pT}} \sum_{v \in \Lambda_q} (1 - \gamma)^{-1} \tilde{f}_K^L(x + u) \mathbf{1}_{\Delta_q}(x + v) \\ &\geq \frac{2\alpha}{|\Lambda_{pT}|} \sum_{u \in \Lambda_{pT}} \tilde{f}_K^L(x + u) \\ &\geq \alpha \mathbb{E} f_K^L \geq \alpha \end{aligned}$$

using (6.7) for the first inequality and Lemma 5.1 for the second. \square

We also define an additional exceptional set

$$E_L^2 := \{x \in \Delta_q : \exists 0 < y \leq \psi(AQ_{x,L}) \text{ such that } x + y \notin \Delta_q\}. \quad (6.9)$$

Note that $E_L^2 \subset \{\Delta_q \cap (-\Lambda_q)^\gamma\} \cup \{-\Lambda_q + (-\psi(AT), 0]\}$. By our choice of q , we see that

$$\mathbb{P}(E_L^2) \leq \delta/8 + |\Lambda_q| \psi(AT) \leq \frac{\delta}{4}.$$

7 Restricting a Family to $(-\Lambda_q)^\gamma$

As noted in the heuristic outline, if we wish to change Q_x on the set $(-\Lambda_q)^\gamma$, we must change f_1^L, \dots, f_{K-1}^L and X_1^L, \dots, X_{K-1}^L as well, since these need no longer satisfy (4) with the new value of Q_x . In order to find suitable replacement functions on $(-\Lambda_q)^\gamma$, we will use take a Step $(K-1, M, M)$ family $\{S, R', g_1, \dots, g_{K-1}, Z_1, \dots, Z_{K-1}, E', Q'_x\}$ with suitable parameters, and then restrict the functions g_h to the set $(0, \gamma s_q) + q\mathbb{Z}$ (multiplying them by $|\Lambda_q|$ so that their ℓ^1 norm is γ times its previous value). The averages along Λ_q starting at any $x \in (-\Lambda_q)^\gamma$ will sample from this set uniformly often, so that we may be able to preserve property (4) there.

This restriction will not preserve the independence of the Z_h as such, since the averages of the restricted g_h must be 0 off of $(-\Lambda_q)^\gamma$. However, we may define the conditional expectations $\mathbb{E}(X \mid \Sigma)$; and these will remain independent, for $\Sigma \in (-\Lambda_q)^\gamma$.

Definition Let X be a random variable X on a discrete probability space (Ω, \mathbb{P}) , and let $\Sigma \subset \Omega$ with $\mathbb{P}(\Sigma) > 0$. We define the *conditional expectation* $\mathbb{E}(X | \Sigma)$ to be the random variable on Σ with $\mathbb{E}(X | \Sigma)(x) = X(x) \forall x \in \Sigma$, where Σ is equipped with the probability measure $\mathbb{P}_\Sigma(x) = \mathbb{P}(\Sigma)^{-1}\mathbb{P}(x)$.

Now we can state the actual form of this restriction for an entire Step (K, M, L) family:

Lemma 7.1. *Let $\{T, R, f_1, \dots, f_K, X_1, \dots, X_K, E, Q_x\}$ satisfy the properties of Step (K, M, L) for the parameters A, δ, D . Say we have $q, B \in \mathcal{Q}$ with q, B and T pairwise relatively prime, and $qB \leq A$. Let*

$$\begin{aligned}\Xi_q^\gamma &:= (0, \gamma s_q) + q\mathbb{Z} \\ \bar{f}_h(x) &:= |\Lambda_q| f_h(x) \mathbf{1}_{\Xi_q^\gamma}(x) \\ \bar{X}_h(x) &:= X_h(x) \mathbf{1}_{(-\Lambda_q)^\gamma}(x) \\ \bar{E} &:= E \cap (-\Lambda_q)^\gamma \\ \bar{Q}_x &:= qBQ_x \\ \bar{A} &:= \frac{A}{qB} \\ \bar{D} &:= \frac{D}{(D, qB)} \\ \bar{T} &:= qBT.\end{aligned}$$

Then $(\bar{T}, D) = 1$ and

- (1) $\bar{f}_1, \dots, \bar{f}_K \in \ell^1(\mathbb{Z}_{\bar{T}})$ with $\bar{f}_h \geq 0$ and $\mathbb{E}\bar{f}_h \leq \gamma \mathbb{E}f_h$
- (2) $\bar{X}_1, \dots, \bar{X}_K \in \ell^1(\mathbb{Z}_{R\bar{T}})$ such that for any nonempty q -periodic $\Sigma \subset (-\Lambda_q)^\gamma$, $\mathbb{E}(\bar{X}_1 | \Sigma), \dots, \mathbb{E}(\bar{X}_K | \Sigma)$ are independent and $\mathbb{E}(\bar{X}_h | \Sigma) \stackrel{d}{=} X_h$.
- (3) $\bar{E} \subset \mathbb{Z}_{\bar{T}}$ with $\mathbb{P}(\bar{E}) \leq \delta \mathbb{P}((-\Lambda_q)^\gamma)$
- (4) $\bar{Q}_x : \mathbb{Z}_{\bar{T}} \rightarrow \mathbb{N}$ such that $\bar{Q}_x | \bar{T} \forall x \in \mathbb{Z}_{\bar{T}}$ and for each $x \notin \bar{E}$,

$$\frac{1}{|\Lambda_{\bar{Q}_x}|} \sum_{a \in \Lambda_{\bar{Q}_x}} \bar{f}_h(x+a) \geq \bar{X}_h(x) \quad \forall 1 \leq h \leq K$$

and

$$\bar{f}_h(x+y-Q_x) = \bar{f}_h(x+y) \quad \forall 1 \leq h \leq K, \bar{Q}_x \leq y \leq \psi(\bar{A}\bar{Q}_x).$$

Proof. Since $(q, T) = 1$, any q -periodic set contains an equal portion of integers from each residue class modulo T . This fact quickly implies properties (1)-(3), noting that the joint distribution of the X_h on any q -periodic $\Sigma \subset (-\Lambda_q)^\gamma$ is the same as their joint distribution on \mathbb{Z}_{RT} . Now for property (4), take $x \in (-\Lambda_q)^\gamma \setminus \bar{E}$ (since it is trivial otherwise). Since Q_x , q , and B are relatively prime,

$$\begin{aligned}\frac{1}{|\Lambda_{\bar{Q}_x}|} \sum_{a \in \Lambda_{\bar{Q}_x}} \bar{f}_h(x+a) &= \frac{1}{|\Lambda_B|} \sum_{u \in \Lambda_B} \frac{1}{|\Lambda_{qQ_x}|} \sum_{z \in \Lambda_{qQ_x}} \bar{f}_h(x+z) \\ &= \frac{1}{|\Lambda_q|} \sum_{v \in \Lambda_q} |\Lambda_q| \mathbf{1}_{\Xi_q^\gamma}(x+v) \cdot \frac{1}{|\Lambda_{Q_x}|} \sum_{w \in \Lambda_{Q_x}} f_h(x+w) \\ &\geq \frac{1}{|\Lambda_{Q_x}|} \sum_{w \in \Lambda_{Q_x}} f_h(x+w) \geq X_h(x) = \bar{X}_h(x)\end{aligned}$$

since if $x \in (-\Lambda_q)^\gamma$, there must exist some $v \in \Lambda_q$ such that $x+v \in \Xi_q^\gamma$.

Finally, $\bar{A}\bar{Q}_x = A\bar{Q}_x$ so (4.6) implies the last claim trivially. \square

8 Completion of the Inductive Step

Now we are ready to define the other functions and prove Step $(K, M, L + 1)$; but since so much goes into this step, we will show the origins of the various pieces.

We began with $\{T_L, R_L, f_1^L, \dots, f_K^L, X_1^L, \dots, X_K^L, E_L, Q_{x,L}\}$ satisfying **(1)-(4)** with the parameters $A, \delta/4$, and D . We modify these objects in several ways, using a new $p = p_{i(K,L)}$ and $q = q_{j(K,L)}$ chosen in Section 6.

First, we applied the p -periodic rearrangement $f \rightarrow \tilde{f}^\omega$ defined in (5.1) to the functions f_h^L and X_h^L , with ω chosen as in Lemma 5.1, and defined an associated exceptional set $E_L^1(\omega)$ in (5.2). We defined sets $\Delta_q, \Psi_q \subset \mathbb{Z}_q$ satisfying (6.2)-(6.7), then we defined the function f_K^{L+1} in (6.8) and an additional exceptional set E_L^2 in (6.9).

We are now ready to proceed.

Set $A_L := AT_Lpq$ and $D_L := DT_Lpq$. By our strong inductive hypothesis, we may assume Step $(K-1, M, M)$ and thus construct a family $\{S, R', g_1, \dots, g_{K-1}, Z_1, \dots, Z_{K-1}, E', Q'_x\}$ satisfying **(1)-(4)** with parameters $A_L, \delta/4, D_L$. (In particular, S and R' are each relatively prime to T_L, p, q , and D .)

Applying Lemma 7.1 with $B = Tp$, we obtain $\{\bar{S}, R', \bar{g}_1, \dots, \bar{g}_{K-1}, \bar{Z}_1, \dots, \bar{Z}_{K-1}, \bar{E}', \bar{Q}'_x\}$ on $(-\Lambda_q)^\gamma$ satisfying **(1)-(4)** with the parameters $\bar{A}_L = A, \delta/4$, and $\bar{D}_L = D$.

Then we define

$$T_{L+1} := \bar{S} = ST_Lpq \quad (8.1)$$

$$E_{L+1} := E_L^1(\omega) \cup E_L^2 \cup \bar{E}' \quad (8.2)$$

$$Q_{x,L+1} := \begin{cases} \tilde{Q}_{x,L}, & x \in \Delta_q \\ \bar{Q}'_x, & x \notin \Delta_q, \end{cases} \quad (8.3)$$

and for $1 \leq h \leq K-1$,

$$f_h^{L+1} := \tilde{f}_h^L \mathbf{1}_{\Delta_q} + \bar{g}_h \mathbf{1}_{\mathbb{Z} \setminus \Delta_q} \in \ell^1(\mathbb{Z}_{T_{L+1}}) \quad (8.4)$$

$$X_h^{L+1} := \tilde{X}_h^L \mathbf{1}_{\Delta_q} + \bar{Z}_h \mathbf{1}_{\mathbb{Z} \setminus \Delta_q} \in \ell^1(\mathbb{Z}_{T_{L+1}}). \quad (8.5)$$

We have already defined

$$f_K^{L+1} := (1 - \gamma)^{-1} \tilde{f}_K^L \mathbf{1}_{\Delta_q} \in \ell^1(\mathbb{Z}_{T_{L+1}}).$$

It thus remains to define X_K^{L+1} . As noted before, we cannot simply define it as in (2.3); we must reduce it slightly so that it equals $Y_{L+1,\gamma,\alpha}$ in distribution.

Recall that $\mathbb{P}(\Delta_q) \geq 1 - \gamma$ and $\mathbb{P}(\Psi_q) > \alpha\gamma$; we have taken γ, α to be dyadic rationals and assumed (6.6), so we may write

$$\frac{1 - \gamma}{\mathbb{P}(\Delta_q)} = \frac{s}{r}, \quad \frac{\alpha\gamma}{\mathbb{P}(\Psi_q)} = \frac{t}{r} \quad (8.6)$$

with $(r, D) = 1$. (Recall that D is odd.) Everything so far is periodic with period $T_{L+1}R_LR'$, so if we define

$$\begin{aligned} R_{L+1} &:= R_L R' r \\ \Gamma_s &:= [0, sR' R_L T_{L+1}) + R_{L+1} T_{L+1} \mathbb{Z} \\ \Gamma_t &:= [0, tR' R_L T_{L+1}) + R_{L+1} T_{L+1} \mathbb{Z} \\ X_K^{L+1} &:= (1 - \gamma)^{-1} \tilde{X}_K^L \mathbf{1}_{\Delta_q} \mathbf{1}_{\Gamma_s} + \alpha \mathbf{1}_{\Psi_q} \mathbf{1}_{\Gamma_t}, \end{aligned}$$

then $X_K^{L+1} \stackrel{d}{=} Y_{L+1,\gamma,\alpha}$ and all of the X_h^{L+1} are periodic with period $R_{L+1}T_{L+1}$.

We now have a family $\{T_{L+1}, R_{L+1}, f_1^{L+1}, \dots, f_K^{L+1}, X_1^{L+1}, \dots, X_K^{L+1}, E_{L+1}, Q_{x,L+1}\}$, with $T_{L+1} \in \mathcal{Q}$ and $(T_{L+1}, D) = (R_{L+1}, D) = 1$. We must check the four properties of Step $(K, M, L+1)$:

(1) $f_1^{L+1}, \dots, f_K^{L+1} \in \ell^1(\mathbb{Z}_{T_{L+1}})$ with $f_h^{L+1} \geq 0$ and $1 \leq \mathbb{E}f_h^{L+1} \leq (1+4\varepsilon)^{(K-1)M+L+1}$.

Proof. For $1 \leq h \leq K-1$, by the inductive hypothesis we see

$$\begin{aligned} \mathbb{E}f_h^{L+1} &= \mathbb{E}\tilde{f}_h^L \mathbb{P}(\Delta_q) + \mathbb{E}\bar{g}_h^L \\ &\leq (1-\gamma+\varepsilon)\mathbb{E}f_h^L + \gamma\mathbb{E}g_h^L \\ &\leq (1+\varepsilon)(1+4\varepsilon)^{(K-1)M+L} \\ &\leq (1+4\varepsilon)^{(K-1)M+L+1}. \end{aligned}$$

For f_K^{L+1} , this follows from (6.8) and the inductive hypothesis. \square

(2) $X_1^{L+1}, \dots, X_K^{L+1} \in \ell^1(\mathbb{Z}_{R_{L+1}T_{L+1}})$ pairwise independent with $X_h^{L+1} \stackrel{d}{=} Y_{M,\gamma,\alpha}$ for $1 \leq h \leq K-1$, and $X_K^{L+1} \stackrel{d}{=} Y_{L+1,\gamma,\alpha}$.

Proof. We have inductively assumed that X_1^L, \dots, X_K^L are pairwise independent; since the periodic rearrangement preserves joint distribution, this is true of $\tilde{X}_1^L, \dots, \tilde{X}_K^L$ as well. We begin by considering the conditional expectations of the \tilde{X}_h^L on Δ_q . Recall that q is relatively prime to p, T_L, R_L, S , and R' , so that Δ_q contains an equal proportion of all residue classes modulo pT_LR_LSR' . Furthermore, Γ_s and Γ_t each contain an equal proportion of all residue classes modulo $R'R_LT_{L+1}$. Thus for any $h < K$ and $\lambda_h, \lambda_K > 0$,

$$\begin{aligned} \mathbb{P}(x \in \Delta_q : X_h^{L+1}(x) \geq \lambda_h, X_K^{L+1}(x) \geq \lambda_K) &= \mathbb{P}(x \in \Delta_q \cap \Gamma_s, \tilde{X}_h^L(x) \geq \lambda_h, \tilde{X}_K^L(x) \geq \lambda_K) \\ &= \frac{s}{r} \mathbb{P}(\Delta_q) \mathbb{P}(\tilde{X}_h^L(x) \geq \lambda_h) \mathbb{P}(\tilde{X}_K^L(x) \geq \lambda_K) \\ &= \mathbb{P}(\Delta_q) \mathbb{P}(X_h^{L+1}(x) \geq \lambda_h) \mathbb{P}(X_K^{L+1}(x) \geq \lambda_K). \end{aligned}$$

Thus $\mathbb{E}(X_h^{L+1} \mid \Delta_q)$ and $\mathbb{E}(X_K^{L+1} \mid \Delta_q)$ are independent; similarly, $\mathbb{E}(X_h^{L+1} \mid \Delta_q)$ and $\mathbb{E}(X_{h'}^{L+1} \mid \Delta_q)$ are independent for any $h < h' < K$.

We proceed similarly on the rest of \mathbb{Z} , letting Σ denote either Ψ_q or $\mathbb{Z}_{T_{L+1}} \setminus (\Delta_q \cup \Psi_q)$; on each of these, Lemma 7.1 implies that $\mathbb{E}(\tilde{Z}_1^L \mid \Sigma), \dots, \mathbb{E}(\tilde{Z}_{K-1}^L \mid \Sigma)$ are pairwise independent and distributed like $Y_{M,\gamma,\alpha}$. Clearly $\mathbb{E}(\mathbf{1}_{\Gamma_t} \mid \Sigma)$ is independent of any of these, so $\mathbb{E}(X_1^{L+1} \mid \Sigma), \dots, \mathbb{E}(X_K^{L+1} \mid \Sigma)$ are pairwise independent for $\Sigma = \Psi_q$ or $\mathbb{Z}_{T_{L+1}} \setminus (\Delta_q \cup \Psi_q)$.

Now for each $1 \leq h < K$,

$$\mathbb{E}(X_h^{L+1} \mid \Delta_q) \stackrel{d}{=} \mathbb{E}(X_h^{L+1} \mid \Psi_q) \stackrel{d}{=} \mathbb{E}(X_h^{L+1} \mid \mathbb{Z}_{T_{L+1}} \setminus (\Delta_q \cup \Psi_q)) \stackrel{d}{=} X_h^{L+1} \stackrel{d}{=} Y_{M,\gamma,\alpha}, \quad (8.7)$$

and so for any $1 \leq h < h' \leq K$ and $\lambda_h, \lambda_{h'} > 0$,

$$\begin{aligned}
\mathbb{P}(X_h^{L+1}(x) \geq \lambda_h, X_{h'}^{L+1}(x) \geq \lambda_{h'}) &= \sum_{\Sigma} \mathbb{P}(x \in \Sigma : X_h^{L+1}(x) \geq \lambda_h, X_{h'}^{L+1}(x) \geq \lambda_{h'}) \\
&= \sum_{\Sigma} \mathbb{P}(\mathbb{E}(X_h^{L+1} | \Sigma)(x) \geq \lambda_h, \mathbb{E}(X_{h'}^{L+1} | \Sigma)(x) \geq \lambda_{h'}) \mathbb{P}(\Sigma) \\
&= \sum_{\Sigma} \mathbb{P}(\mathbb{E}(X_h^{L+1} | \Sigma)(x) \geq \lambda_h) \mathbb{P}(\mathbb{E}(X_{h'}^{L+1} | \Sigma)(x) \geq \lambda_{h'}) \mathbb{P}(\Sigma) \\
&= \sum_{\Sigma} \mathbb{P}(X_h^{L+1}(x) \geq \lambda_h) \mathbb{P}(\mathbb{E}(X_{h'}^{L+1} | \Sigma)(x) \geq \lambda_{h'}) \mathbb{P}(\Sigma) \\
&= \mathbb{P}(X_h^{L+1}(x) \geq \lambda_h) \sum_{\Sigma} \mathbb{P}(x \in \Sigma : X_{h'}^{L+1}(x) \geq \lambda_{h'}) \\
&= \mathbb{P}(X_h^{L+1}(x) \geq \lambda_h) \mathbb{P}(X_{h'}^{L+1}(x) \geq \lambda_{h'}).
\end{aligned}$$

(The sum is over the sets $\Sigma = \Delta_q, \Psi_q$, and $\mathbb{Z}_{T_{L+1}} \setminus (\Delta_q \cup \Psi_q)$; the property (8.7) enters in at the fourth equality.) Thus we have preserved independence.

We have already noted that $X_K^{L+1} \stackrel{d}{=} Y_{L+1, \gamma, \alpha}$. □

(3) An exceptional set $E_{L+1} \subset \mathbb{Z}_{T_{L+1}}$ with $\mathbb{P}(E_{L+1}) \leq \delta$.

Proof. $\mathbb{P}(E_{L+1}) \leq \mathbb{P}(E_L^1(\omega)) + \mathbb{P}(E_L^2) + \mathbb{P}(E_L') \leq \delta/2 + \delta/4 + \delta\gamma/4 \leq \delta$. □

(4) $Q_x = Q_{x, L+1} : \mathbb{Z}_{T_{L+1}} \rightarrow \mathbb{Z}^+$ with $Q_x \in \mathcal{Q}$ and $Q_x \mid T_{L+1} \forall x$, such that for each $x \notin E_{L+1}$,

$$\frac{1}{|\Lambda_{Q_x}|} \sum_{a \in \Lambda_{Q_x}} f_h^{L+1}(x+a) \geq X_h^{L+1}(x) \quad \forall 1 \leq h \leq K$$

and

$$f_h(x+y-Q_x) = f_h(x+y) \quad \forall 1 \leq h \leq K, Q_x \leq y \leq \psi(AQ_x).$$

Proof. Since $\tilde{Q}_{x, L} \mid T_L$ and $\bar{Q}'_x \mid \bar{S}$, clearly $Q_{x, L+1} \mid T_{L+1}$; in addition, $\tilde{Q}_{x, L} \in \mathcal{Q}$ and $\bar{Q}'_x = pqT_LQ'_x \in \mathcal{Q}$ (this is squarefree since p, q, T_L , and Q'_x are pairwise relatively prime).

For $x \in \Delta_q \setminus E_{L+1}$ and $1 \leq y \leq \psi(AT_L)$, we see that

$$f_h^{L+1}(x+y) = \tilde{f}_h^L(x+y) = f_h^L(x+y + \xi_i(\omega))$$

for some ω fixed and i depending only on x . Thus (4) follows from the previous step.

On Ψ_q and $\mathbb{Z} \setminus (\Delta_q \cup \Psi_q)$, since $pqT_L \mid \bar{Q}'_x$, this is just Lemma 7.1 for $h < K$, and for $h = K$ this is Lemma 6.1 combined with the observation that f_K^{L+1} is periodic by pqT_L . □

Thus we have proved Step $(K, M, L+1)$; by induction, we have proved Proposition 4.1 and Theorem 3.1.

9 Notes on the Proof

Several of the conditions, parameters and lemmas in this complicated argument appear on a first reading to be extraneous to the proof. In the interest of clarity, we find it helpful to outline in hindsight the purposes of the following:

- The condition (3.3) lets us prove Theorem 3.1 from Proposition 4.1 (note that this proof requires $\gamma/C\varepsilon_\gamma \rightarrow \infty$ as $\gamma \rightarrow 0$). (3.4) comes in at (6.1) and the subsequent definition of Ψ_q , and (3.5) allows us to claim (4.7).
- Prescribing an exact distribution $Y_{n,\gamma,\alpha}$ (defined in (4.4)) for the X_h^L is necessary in order to guarantee (8.7), which ensures that pairwise independence is preserved.
- γ and α must be dyadic rationals, and Δ_q and Ψ_q must be chosen to satisfy (6.6), so that we can assume $(r, D) = 1$ in (8.6), so that we can have $(R, D) = 1$ in Step (K, M, L) , so that we can choose our $(K-1, M, M)$ family with $(R', q) = 1$, so that X_K^{L+1} will be independent of the other X_h^{L+1} .
- The parameter D lets us guarantee that when we inductively introduce a Step $(K-1, M, M)$ family, we can ensure that its period $R'S$ is relatively prime to T_L, p , and q , thus allowing us to apply Lemma 7.1.
- We have two distinct parameters T and R because we will need the period of f_K^L to be squarefree in Lemma 6.1 (because the result of [7] only applies for squarefree moduli), but the operation of reducing X_K^{L+1} to its proper distribution will multiply its period by a large power of 2.
- The condition (4.6) is necessary for (4.7), connecting the actual averages over the sequence $\{n_k\}$ with the averages over a set of residues $|\Lambda_Q| \in \mathbb{Z}_Q$. The parameter A must be allowed to take arbitrarily large values, although it need only be ≥ 1 when used in (4.7), because each application of Lemma 7.1 divides it by a large constant.
- The periodic rearrangement defined in (5.1) puts a uniform lower bound on the averages of \tilde{f}_K^L over $Q_{x,L+1}$ on a set which depends only on q and not on f_K^L ; this allows us to prove Lemma 6.1.

The author thanks his dissertation advisor, M. Christ, for several key suggestions including the external randomization, and M. Wierdl for many encouraging and stimulating discussions on this topic.

References

- [1] J. Bourgain. An approach to pointwise ergodic theorems. In *Geometric aspects of functional analysis (1986/87)*, volume 1317 of *Lecture Notes in Math.*, pages 204–223. Springer, Berlin, 1988.
- [2] Jean Bourgain. Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Études Sci. Publ. Math.*, (69):5–45, 1989. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.
- [3] Zoltán Buczolich and R. Daniel Mauldin. Divergent square averages. <http://arxiv.org/abs/math/0504067v2>.
- [4] Zoltán Buczolich and R. Daniel Mauldin. Concepts behind divergent ergodic averages along the squares. In *Ergodic theory and related fields*, volume 430 of *Contemp. Math.*, pages 41–56. Amer. Math. Soc., Providence, RI, 2007.
- [5] Stephen D. Cohen. Uniform distribution of polynomials over finite fields. *J. London Math. Soc. (2)*, 6:93–102, 1972.

- [6] Jean-Pierre Conze. Convergence des moyennes ergodiques pour des sous-suites. In *Contributions au calcul des probabilités*, pages 7–15. Bull. Soc. Math. France, Mém. No. 35. Soc. Math. France, Paris, 1973.
- [7] Andrew Granville and Pär Kurlberg. Poisson statistics via the Chinese remainder theorem. *Adv. Math.*, 218(6):2013–2042, 2008.
- [8] Christopher Hooley. On the difference between consecutive numbers prime to n . II. *Publ. Math. Debrecen*, 12:39–49, 1965.
- [9] S. Sawyer. Maximal inequalities of weak type. *Ann. of Math. (2)*, 84:157–174, 1966.
- [10] Terence Tao and Van Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [11] Máté Wierdl. Pointwise ergodic theorem along the prime numbers. *Israel J. Math.*, 64(3):315–336 (1989), 1988.

PATRICK LAVICTOIRE
 DEPARTMENT OF MATHEMATICS, UC BERKELEY
 BERKELEY, CA 94720-3840 USA
E-mail: patlavic@math.berkeley.edu
URL: <http://math.berkeley.edu/~patlavic/>